6-2013

An Analysis of Polynomials that Commute Under Composition

Samuel J. Williams
Union College - Schenectady, NY

An Analysis of Polynomials that Commute Under Composition

By

Samuel Williams

* * * * * * * * *

Submitted in partial fulfillment

of the requirements for

Honors in the Department of Mathematics

UNION COLLEGE

June, 2013

ABSTRACT

WILLIAMS, SAMUEL An Analysis of Polynomials that Commute Under Composition
Department of Mathematics, June 2013.

ADVISOR: Professor Karl Zimmermann

It is well known that polynomials commute under addition and multiplication. It turns out that certain polynomials also commute under composition. In this paper, we examine polynomials with coefficients in the field of complex numbers that commute under composition (also referred to as "commuting polynomials"). We begin this examination by defining what it means for polynomials to commute under composition. We then introduce sequences of commuting polynomials and observe how the polynomials in these sequences (later defined as chains) along with other commuting polynomials relate to a concept called similarity. These observations allow us to better understand the qualities and characteristics of polynomials that commute under composition. Later, we characterize all chains of polynomials with complex coefficients. We conclude the paper with an exploration of a concept called self similarity and how it relates to commuting polynomials. This exploration reinforces the observations we made about the qualities and characteristics of commuting polynomials and also provides new insights. The examination outlined in this paper provides us with an accessible understanding of polynomials that commute under composition.

Table of Contents

1 Introduction

It is well known that polynomials commute under addition and multiplication. However, polynomials do not necessarily commute under composition, though some polynomials do.

Let **C** denote the field of complex numbers. Recall that $f(x)$ is a polynomial if and only if it is of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_2 x^2 + a_1 x + a_0,$$

where, for this paper, the coefficients, $a_1, \ ..., \ a_n \in \mathbf{C}$. That is, all polynomials discussed in this paper are in $\mathbf{C}[x]$, the set of polynomials with complex coefficients. Before we begin, we will first recall the following definitions.

Definition 1.1 *Let*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_2 x^2 + a_1 x + a_0$$

be a polynomial. If $a_n \neq 0$*, then we say that the degree of f(x) is n. In other words, the degree is the highest exponent of any x with a corresponding, non-zero coefficient.*

Now recall that, for any $f(x)$, $g(x) \in \mathbf{C}[x]$, there exists $q(x)$, the *composition* of $f(x)$ with $g(x)$, written

$$q(x) = (f \circ g)(x).$$

The resulting polynomial $q(x)$ is simply

$$q(x) = f(g(x)).$$

In other words, the composition of f and g is just the result of plugging $g(x)$ into $f(x)$.

Definition 1.2 *Assume f(x) and g(x) are polynomials. We say $f(x)$ and $g(x)$ commute under composition if and only if*

$$(f \circ g)(x) = (g \circ f)(x).$$

This paper is about commuting polynomials. We will examine these polynomials and their properties. In addition, we will characterize certain sequences of polynomials. From here on out, "commute" will be taken to mean "commute under composition."

2 Commuting Polynomials and Similarity

We begin with examples of polynomials that commute under composition. Consider the polynomials

$$f(x) = x$$

and

$$g(x) = x^2 + 1.$$

We can show that

$$(f \circ g)(x) = (g \circ f)(x)$$

as

$$f(g(x)) = (x^2 + 1) = (x)^2 + 1 = g(f(x)).$$

This is a basic example that illustrates a dramatic result: $f(x) = x$ commutes with *every* polynomial.

Proposition 2.1 *The polynomial $f(x) = x$ commutes with any polynomial $g(x) \in \mathbf{C}[x]$.*

Proof. Take any polynomial $g(x) \in \mathbf{C}[x]$ and let $f(x) = x$. Now, $(f \circ g)(x)$ is given by

$$f(g(x)) = g(x) = g(f(x)).$$

Thus, $(f \circ g)(x) = (g \circ f)(x)$, as desired. □

In addition to simple examples, there are also more complicated examples of commuting polynomials.

Definition 2.2 *The Chebyshev polynomials of the first kind,* $T_n(x)$ *with* $n \geq 1$*, are given by*

$$T_n(x) = \cos n(\cos^{-1}(x)).$$

(Chebyshev is spelled Tchebychef by Barbeau [1], from whom we get the above notation). The second Chebyshev polynomial, $T_2(x) = \cos 2(\cos^{-1}(x))$, can be expressed using the polynomial structure mentioned at the beginning of this paper by recalling the well known trigonometry identity, $\cos(2u) = 2\cos^2(u) - 1$.

$$\begin{aligned} T_2(x) &= \cos 2(\cos^{-1}(x)) \\ &= \cos(2u), \text{where} \quad u = \cos^{-1}(x) \\ &= 2\cos^2(u) - 1 \\ &= 2\cos^2(\cos^{-1}(x)) - 1 \\ &= 2x^2 - 1. \end{aligned}$$

Similar evaluations can be performed on the other Chebyshev polynomials, which shows that

$$\begin{aligned} T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ &\cdot \\ &\cdot \end{aligned}$$

.

It turns out that that higher degree Chebyshev polynomials can be related to lower degree Chebyshev polynomials. In fact, we can express a given Chebyshev polynomial in terms of Chebyshev polynomials of a lower degree, again using trigonometry. First, note that

$$\begin{aligned}\cos(n-2)x &= \cos((n-1)x+(-x)) \\ &= \cos(-x)\cos(n-1)x - \sin(-x)\sin(n-1)x \\ &= \cos(x)\cos((n-1)x) + \sin(x)\sin((n-1)x).\end{aligned}$$

When we rearrange the equation, we have

$$\sin(x)\sin((n-1)x) = \cos((n-2)x) - \cos(x)\cos((n-1)x). \quad (*)$$

Now, we calculate

$$\begin{aligned}T_n(x) &= \cos n(\cos^{-1}(x)) \\ &= \cos n(u), \quad \text{where} \quad u = \cos^{-1}(x) \\ &= \cos(nu) \\ &= \cos(u + nu - u) \\ &= \cos(u + (nu - u)) \\ &= \cos(u + (n-1)u) \\ &= \cos(u)\ cos((n-1)u) - \sin(u)\sin((n-1)u)\end{aligned}$$

$$
\begin{aligned}
&= \cos(u)\cos((n-1)u) - (\cos((n-2)u) - \cos(u)\cos((n-1)u)) \quad \text{by} \quad \text{equation} \quad (*) \\
&= \cos(u)\cos((n-1)u) - \cos((n-2)u) + \cos(u)\cos((n-1)u) \\
&= 2\cos(u)\cos((n-1)u) - \cos((n-2)u).
\end{aligned}
$$

When we replace u with $\cos^{-1}(x)$, we have

$$T_n(x) = 2\cos(\cos^{-1}(x))\cos((n-1)(\cos^{-1}(x))) - \cos((n-2)(\cos^{-1}(x)))$$

or, by definition

$$T_n = 2xT_{n-1}(x) - T_{n-2}(x).$$

We have already shown that $T_1(x)$ commutes with any $T_n(x)$, as $T_1(x) = x$. We will also show that other Chebyshev polynomials commute with one another. Before we move on to this general result, we will show that two low degree Chebyshev polynomials, $T_2(x)$ and $T_3(x)$, commute with one another.

Proposition 2.3 *The Chebyshev polynomials $T_2(x)$ and $T_3(x)$ commute with one another.*

Proof. To show that $T_2(x)$ and $T_3(x)$ commute with one another, we simply calculate

$$
\begin{aligned}
T_2(T_3(x)) &= 2(4x^3 - 3x)^2 - 1 \\
&= 2(16x^6 - 24x^4 + 9x^2) - 1 \\
&= 32x^6 - 48x^4 + 18x^2 - 1 \\
&= (32x^6 - 48x^4 + 24x^2 - 4) - (6x^2 - 3)
\end{aligned}
$$

$$
\begin{aligned}
&= 4(8x^6 - 6x^4 + 4x^2 - 1) - 3(2x^2 - 1) \\
&= 4(2x^2 - 1)^3 - 3(2x^2 - 1) \\
&= T_3(T_2(x)).
\end{aligned}
$$

Thus, we conclude that

$$(T_2 \circ T_3)(x) = (T_3 \circ T_2)(x),$$

as desired. □

The above proof involves the evaluation of polynomials. While this is an approachable method for the lower degree Chebyshev Polynomials, calculating the coefficients of $T_{43}(x)$ and $T_{88}(x)$ would be time consuming and difficult. Therefore, it behooves us to establish that any Chebyshev polynomial commutes with any other Chebyshev polynomial via a different form of analysis.

Proposition 2.4 *For any Chebyshev polynomials $T_n(x)$ and $T_m(x)$, $T_n(x)$ commutes with $T_m(x)$ under composition.*

Proof. Assume that $T_n(x)$ and $T_m(x)$ are Chebyshev polynomials. Then

$$T_n(x) = \cos n(\cos^{-1}(x))$$

and

$$T_m(x) = \cos m(\cos^{-1}(x))$$

by definition. Then $(T_n \circ T_m)(x)$ is

$$\begin{aligned}
T_n(T_m(x)) &= \cos n(\cos^{-1}(T_m(x))) \\
&= \cos n(\cos^{-1}(\cos m(\cos^{-1}(x)))) \\
&= \cos n(m(\cos^{-1}(x))) \\
&= \cos(nm \cos^{-1}(x)) \\
&= \cos(mn \cos^{-1}(x)) \\
&= \cos m(n(\cos^{-1}(x))) \\
&= \cos m(\cos^{-1}(\cos n(\cos^{-1}(x)))) \\
&= \cos m(\cos^{-1}(T_n(x))) \\
&= T_m(T_n(x)).
\end{aligned}$$

Thus, $(T_n \circ T_m)(x) = (T_m \circ T_n)(x)$, as desired. □

It is easy to see that the Chebyshev polynomials commute with one another by examining the trig definition. That is, a given Chebyshev polynomial commutes with every Chebyshev polynomial. Let us now examine another set of polynomials.

Consider the set of polynomials defined by $\{P_n(x) = x^n,\ \ n \in \mathbf{N}\}$, where P stands for power. Obviously, $P_1(x) = x$ is a power polynomial, and $P_1(x)$ commutes with every polynomial with complex coefficients. So, we have that the first power polynomial commutes with every polynomial (with complex coefficients) under composition. While we cannot say the same for the rest of the power polynomials, we can show that the power polynomials, like the Chebyshev polynomials, commute with one another under composition.

Proposition 2.5 *Assume $P_n(x)$ and $P_m(x)$ are power polynomials. Then $P_n(x)$ and $P_m(x)$ commute with one another under composition.*

Proof. We compute $(P_n \circ P_m)(x)$ to be

$$\begin{aligned} P_n(P_m(x)) &= (P_m(x))^n \\ &= (x^m)^n \\ &= x^{mn} \\ &= x^{nm} \\ &= (x^n)^m \\ &= (P_n(x))^m \\ &= P_m(P_n(x)). \end{aligned}$$

It follows that $(P_n \circ P_m)(x) = (P_m \circ P_n)(x)$, as desired. □

We will soon introduce a definition that will allow us to create new examples of commuting polynomials. First, recall that any degree one polynomial $f(x) = ax + b$ is invertible. The inverse polynomial is given by $f^{-1}(x) = \dfrac{x-b}{a}$. We know that $a \neq 0$ since $f(x)$ is a degree one polynomial. We can now move on to the definition.

Definition 2.6 *Let $f(x)$, $g(x) \in \mathbf{C}[x]$. Then $f(x)$ is similar to $g(x)$ if and* *only if there exists*

$$\lambda(x) = ax + b, a \neq 0$$

such that

$$(\lambda^{-1} \circ f \circ \lambda)(x) = g(x).$$

Such a polynomial is called a similarity.

For example, we can see that $f(x) = x^2$ is similar to $g(x) = x^2 + 2x$ via the similarity $\lambda(x) = x + 1$. To see this note that, $(\lambda^{-1} \circ f \circ \lambda)(x)$ is given by

$$\begin{aligned} \lambda^{-1}(f(\lambda(x)))) &= \lambda^{-1}(f(x+1)) \\ &= \lambda^{-1}((x+1)^2) \\ &= \lambda^{-1}(x^2 + 2x + 1) \\ &= (x^2 + 2x + 1) - 1 \\ &= x^2 + 2x \\ &= g(x) \end{aligned}$$

Therefore, $f(x)$ is similar to $g(x)$, as desired.

We can show that similarity is an equivalence relation. This result (more specifically: the reflexivity of similarity) will help us obtain results that appear later in the paper.

Theorem 2.7 *Similarity is an equivalence relation.*

Proof. In order to show that similarity is an equivalence relation, we must show that similarity is reflexive, symmetric, and transitive.

To show that similarity is reflexive, let $f(x)$ be any polynomial, and consider the degree 1 polynomial

$$\lambda(x) = x.$$

The inverse of $\lambda(x)$ is given by

$$\lambda^{-1}(x) = x.$$

Then,

$$(\lambda^{-1} \circ f \circ \lambda)(x) = \lambda^{-1}(f(\lambda(x))) = \lambda^{-1}(f(x)) = f(x).$$

So $f(x)$ is similar to $f(x)$ via $\lambda(x)$, and similarity is reflexive, as desired.

To see that similarity is symmetric, take any two polynomials $f(x)$ and $g(x)$, and assume $f(x)$ is similar to $g(x)$ via a similarity $\lambda(x)$. Then

$$(\lambda^{-1} \circ f \circ \lambda)(x) = g(x)$$

by the definition of similarity. Now

$$(\lambda \circ g \circ \lambda^{-1})(x) = (\lambda \circ \lambda^{-1} \circ f \circ \lambda \circ \lambda^{-1})(x) = f(x).$$

But this means that

$$((\lambda^{-1})^{-1} \circ g \circ \lambda^{-1})(x) = f(x)$$

and we have that $g(x)$ is similar to $f(x)$, as $\lambda^{-1}(x)$ is an invertible, degree 1 polynomial. Because $f(x)$ being similar to $g(x)$ implies that $g(x)$ is similar to

$f(x)$, similarity is symmetric, as desired.

To show that similarity is transitive, take any three polynomials $f(x)$, $g(x)$, and $q(x)$. Assume $f(x)$ is similar to $g(x)$ and $g(x)$ is similar to $q(x)$. Then, by definition,

$$(\lambda^{-1} \circ f \circ \lambda)(x) = g(x) \quad \text{where} \quad \lambda(x) \quad \text{is} \quad \text{a} \quad \text{similarity}$$

and

$$(\alpha^{-1} \circ g \circ \alpha)(x) = q(x) \quad \text{where} \quad \alpha(x) \quad \text{is} \quad \text{a} \quad \text{similarity.}$$

It follows that

$$(\alpha^{-1} \circ (\lambda^{-1} \circ f \circ \lambda) \circ \alpha)(x) = q(x)$$

which may be rewritten as

$$((\alpha^{-1} \circ \lambda^{-1}) \circ f \circ (\lambda \circ \alpha))(x) = ((\lambda \circ \alpha)^{-1} \circ f \circ (\lambda \circ \alpha)) = q(x). \quad (*)$$

$\lambda(x)$ and $\alpha(x)$ are similarities so, without loss of generality, we can set

$$\lambda(x) = ax + b$$

and

$$\alpha(x) = cx + d.$$

Therefore,

$$(\lambda \circ \alpha)(x) = a(cx + d) + b = acx + (ad + d), \quad ac \neq 0.$$

So, $(\lambda \circ \alpha)(x)$ satisfies the definition of a similarity. Then, by equation $(*)$, $f(x)$ is similar to $q(x)$. Since $f(x)$ being similar to $g(x)$ and $g(x)$ being similar to $q(x)$ implies that $f(x)$ is similar to $q(x)$, we conclude that similarity is transitive, as desired.

Because similarity is reflexive, symmetric, and transitive, we conclude that similarity is an equivalence relation, as desired. □

We can use similarities to construct new polynomials from given polynomials. Consider the following definition.

Definition 2.8 *let f(x) be a polynomial of the form*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_2 x^2 + a_1 x + a_0, \quad a_n \neq 0.$$

We call a_n the leading coefficient of $f(x)$ and say $f(x)$ is monic if and only if its leading coefficient is 1.

As it turns out, we can show that any polynomial is similar to a monic polynomial. Before we prove this fact, we first recall that the degree of a polynomial can change when it is multiplied by another. For any two polynomials $f(x)$ and $g(x)$, $\deg(f(x)g(x)) = \deg(f(x)) + \deg(f(x))$. Composition also affects the degree of a polynomial.

Proposition 2.9 *Let $f(x)$ and $g(x)$ be polynomials,* $\deg(f(x)) = n$ *and* $\deg(g(x)) = m$. *Then* $\deg(f \circ g)(x) = \deg(f(x)) \deg(g(x)) = nm$.

Proof. Assume $f(x)$ and $g(x)$ are polynomials with $\deg(f(x)) = n$ and $\deg(g(x)) = m$. Without loss of generality, we may assume that the leading coefficient of $f(x)$ is a_n and that the leading coefficient of $g(x)$ is b_m. Now, we know that the composition of $f(x)$ with $g(x)$ is given by $f(g(x))$. Thus, the leading coefficient of $(f \circ g)(x)$ is found in the term

$$a_n(g(x))^n.$$

It can be easily shown that the leading term of $(f \circ g)(x)$ is

$$a_n(b_m x^m)^n.$$

Since we know that $a_n \neq 0 \neq b_m$ by the above definition, we can conclude that the degree of $(f \circ g)(x) = mn$, as desired. □

We can now move on to the following theorem, which allows for the construction of monic polynomials via similarity.

Theorem 2.10 *Assume f(x) is a polynomial. There exists a similarity $\lambda(x)$ such that*

$$(\lambda^{-1} \circ f \circ \lambda)(x) = g(x)$$

where $g(x)$ is a monic polynomial of the same degree as $f(x)$.

Proof. Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$$

be a polynomial of degree n. Then, for some similarity $\lambda(x)$,

$$\lambda(x) = bx + c,$$

we have

$$
\begin{aligned}
(\lambda^{-1} \circ f \circ \lambda)(x) &= \lambda^{-1}(f(\lambda(x))) \\
&= \lambda^{-1}(f(bx+c)) \\
&= \lambda^{-1}(a_n(bx+c)^n + a_{n-1}(bx+c)^{n-1} + ... + a_1(bx+c)x + a_0) \\
&= (\frac{1}{b})(a_n(bx+c)^n + a_{n-1}(bx+c)^{n-1} + ... + a_1(bx+c)x + a_0) - (\frac{c}{b}) \\
&= g(x).
\end{aligned}
$$

We see that $g(x)$ is of degree n. In order for $g(x)$ to be monic, the leading coefficient of $g(x)$ must be 1. The above equations show that the coefficients of highest degree of $g(x)$ are given by

$$(\frac{1}{b})(a_n(bx+c)^n).$$

When expanded, the above formula shows that the leading coefficient of $g(x)$ is

$$\frac{a_n b^n}{b} = a_n b^{n-1}.$$

Consider $b = (\frac{1}{a_n})^{\frac{1}{n-1}}$ and $c = 0$. Then

$$a_n b^{n-1} = a_n(\frac{1}{a_n}^{\frac{1}{n-1}})^{n-1} = a_n(\frac{1}{a_n}) = \frac{a_n}{a_n} = 1.$$

So there exists a similarity $\lambda(x) = (\frac{1}{a_n})^{\frac{1}{n-1}}x$ that shows that $f(x)$ is similar to a monic polynomial of the same degree, as desired. □

We can now use similarities to construct new pairs of commuting polynomials from polynomials that already commute.

Theorem 2.11 *Let* $f(x)$, $g(x)$, $\lambda(x) \in \mathbf{C}[x]$ *where* $\deg(\lambda(x)) = 1$. *Then* $f(x)$ *and* $g(x)$ *commute under composition if and only if* $(\lambda^{-1} \circ f \circ \lambda)(x)$ *and* $(\lambda^{-1} \circ g \circ \lambda)(x)$ *commute under composition.*

Proof. Let $f(x)$, $g(x)$, and $\lambda(x)$ be a polynomials, and let $\deg(\lambda(x)) = 1$. First, assume $f(x)$ and $g(x)$ commute under composition. Then

$$(f \circ g)(x) = (g \circ f)(x)$$

by definition, and if follows that

$$(\lambda^{-1} \circ f \circ g \circ \lambda)(x) = (\lambda^{-1} \circ g \circ f \circ \lambda)(x).$$

Now, because $(\lambda \circ \lambda^{-1})(x)$ is the identity function, we may insert this composition into the middle of each side of the equation without affecting the equality:

$$(\lambda^{-1} \circ f \circ (\lambda \circ \lambda^{-1}) \circ g \circ \lambda)(x) = (\lambda^{-1} \circ g \circ (\lambda \circ \lambda^{-1}) \circ f \circ \lambda)(x).$$

But this means that

$$((\lambda^{-1} \circ f \circ \lambda) \circ (\lambda^{-1} \circ g \circ \lambda))(x) = ((\lambda^{-1} \circ g \circ \lambda) \circ (\lambda^{-1} \circ f \circ \lambda))(x).$$

as desired.

Now, assume that $(\lambda^{-1} \circ f \circ \lambda)(x)$ and $(\lambda^{-1} \circ g \circ \lambda)(x)$ commute under composition. Then

$$((\lambda^{-1} \circ f \circ \lambda) \circ (\lambda^{-1} \circ g \circ \lambda))(x) = ((\lambda^{-1} \circ g \circ \lambda) \circ (\lambda^{-1} \circ f \circ \lambda))(x).$$

which means that

$$(\lambda^{-1} \circ f \circ (\lambda \circ \lambda^{-1}) \circ g \circ \lambda)(x) = (\lambda^{-1} \circ g \circ (\lambda \circ \lambda^{-1}) \circ f \circ \lambda)(x),$$

or

$$(\lambda^{-1} \circ f \circ g \circ \lambda)(x) = (\lambda^{-1} \circ g \circ f \circ \lambda)(x).$$

We may compose each side of the equation with $\lambda(x)$ and $\lambda^{-1}(x)$ as follows

$$(\lambda \circ (\lambda^{-1} \circ f \circ g \circ \lambda) \circ \lambda^{-1}))(x) = (\lambda \circ (\lambda^{-1} \circ g \circ f \circ \lambda) \circ \lambda^{-1}))(x)$$

which we can rewrite as

$$((\lambda \circ \lambda^{-1}) \circ f \circ g \circ (\lambda \circ \lambda^{-1}))(x) = ((\lambda \circ \lambda^{-1}) \circ g \circ f \circ (\lambda \circ \lambda^{-1}))(x).$$

This simplifies to

$$(f \circ g)(x) = (g \circ f)(x).$$

So $f(x)$ and $g(x)$ commute under composition, as desired. □

The above theorem shows that the commutability of polynomials is preserved by similarity. Consider $f(x) = x^2$ and $g(x) = x^3$ (both of which are power polynomials) and $\lambda(x) = x + 1$. We know that $f(x)$ and $g(x)$ commute under composition. To see that $(\lambda^{-1} \circ f \circ \lambda)(x)$ and $(\lambda^{-1} \circ g \circ \lambda)(x)$ commute under composition, we see that

$$(\lambda^{-1} \circ f \circ \lambda)(x) = (x + 1)^2 - 1 = x^2 + 2x$$

and

$$(\lambda^{-1} \circ g \circ \lambda)(x) = (x + 1)^3 - 1 = x^3 + 3x^2 + 3x.$$

Composing these two polynomials together shows

$$\begin{aligned}
((\lambda^{-1} \circ f \circ \lambda) \circ (\lambda^{-1} \circ g \circ \lambda))(x) &= (x^3 + 3x^2 + 3x)^2 + 2(x^3 + 3x^2 + 3x) \\
&= (x^6 + 6x^5 + 15x^4 + 18x^3 + 9x^2) + (2x^3 + 6x^2 + 6x) \\
&= x^6 + 6x^5 + 15x^4 + 20x^3 + 15x^2 + 6x \\
&= (x^6 + 6x^5 + 12x^4 + 8x^3) + (3x^4 + 12x^3 + 12x^2) + \\
&\quad (3x^2 + 6x) \\
&= (x^2 + 2x)^3 + 3(x^2 + 2x)^2 + 3(x^2 + 2x) \\
&= ((\lambda^{-1} \circ g \circ \lambda) \circ (\lambda^{-1} \circ f \circ \lambda))(x).
\end{aligned}$$

Thus, we see an example in which $(\lambda^{-1} \circ f \circ \lambda)(x)$ and $(\lambda^{-1} \circ g \circ \lambda)(x)$ commute under composition because $f(x)$ and $g(x)$ commute under composition.

3 Properties of Commuting Polynomials

An initial observation that sets up a number of interesting and stronger results involves the number of polynomials of a given degree that can commute with a given degree two polynomial. First, the following lemma shall be proven, which allows us to ascertain the possible leading coefficients of a polynomial that commutes with a given monic polynomial.

Lemma 3.1 *If a polynomial $g(x)$ of degree k commutes with a monic polynomial $h(x)$ of degree n, then the leading coefficient of $g(x)$ is a root of $x^{n-1} - 1$.*

Proof. Assume

$$g(x) = \sum_{m=0}^{k} c_m x^m, \quad c_k \neq 0$$

commutes with monic polynomial $h(x)$ of degree n. Since

$$(h \circ g)(x) = (g \circ h)(x),$$

an analysis of the leading coefficients of each polynomial reveals that

$$c_k (x^n)^k = (c_k x^k)^n,$$

which implies that

$$(c_k) x^{nk} = (c_k)^n x^{nk}.$$

When we compare the degree kn coefficients, we see that

$$c_k = (c_k)^n,$$

which we can rewrite as

$$(c_k)^n - c_k = (c_k)((c_k)^{n-1} - 1) = 0.$$

We know that $c_k \neq 0$, so we may conclude that the leading coefficient c_k is a root of

$$x^{n-1} - 1$$

as desired. □

Now that we can determine the possible leading coefficients of polynomials that commute with monic polynomials, we have the means to prove the following theorem, and the proof is similar to the one presented by Rivlin [4].

Theorem 3.2 *There is at most one polynomial of degree $k \geq 1$ that commutes with a given polynomial $s(x)$ of degree 2.*

Proof. Assume

$$s(x) = a_2x^2 + a_1x + a_0, a_2 \neq 0$$

is a quadratic. To prove the theorem, we will show that there cannot be two distinct polynomials of degree $k \geq 1$ that commute with $s(x)$ under composition. To begin, consider the similarity

$$\lambda(x) = \frac{x}{a_2} - \frac{a_1}{2a_2}$$

which shows that $s(x)$ is similar to

$$\begin{aligned}
(\lambda^{-1} \circ s \circ \lambda)(x) &= \lambda^{-1}(a_2(\frac{x}{a_2} - \frac{a_1}{2a_2})^2 + a_1(\frac{x}{a_2} - \frac{a_1}{2a_2}) + a_0) \\
&= a_2(a_2(\frac{x}{a_2} - \frac{a_1}{2a_2})^2 + a_1(\frac{x}{a_2} - \frac{a_1}{2a_2}) + a_0) + \frac{a_2}{2} \\
&= x^2 + a_0a_2 + \frac{a_1}{2} - \frac{{a_1}^2}{4} \\
&= x^2 + c, \quad \text{where} \quad c = a_0a_2 + \frac{a_1}{2} - \frac{{a_1}^2}{4} \\
&= g(x).
\end{aligned}$$

Now, because $s(x)$ is similar to $g(x) = x^2 + c$, showing that there cannot be two distinct polynomials of degree k that commute with $g(x)$ under composition means that there cannot be two distinct polynomials of degree k that commute with $s(x)$. If there existed such polynomials, applying $\lambda(x)$ to each polynomial would create two distinct polynomials of degree k that commute with $g(x)$ by **2.11**, which contradicts the premise. Thus, showing that there cannot be two distinct polynomials of degree k that commute with $g(x)$ completes the proof.

We continue by assuming, for the sake of contradiction, that $p(x)$ and $q(x)$ are distinct polynomials of degree k that both commute with $g(x)$ under composition. Then

$$(p \circ g)(x) = p(x^2 + c) = p^2(x) + c = (g \circ p)(x)$$

and

$$(q \circ g)(x) = q(x^2 + c) = q^2(x) + c = (g \circ q)(x).$$

Since $p(x)$ and $q(x)$ both commute with $g(x)$, a monic polynomial of degree 2, we know, by **3.1**, that the leading coefficient of each polynomial is a root of $x^{2-1} - 1 = x - 1$. So the leading coefficient of both $p(x)$ and $q(x)$ must be 1. Let

$$r(x) = (p - q)(x).$$

Because the leading coefficient of $p(x)$ and $q(x)$ is 1, we know that $\deg(r(x)) = t \leq k - 1$. When we examine $(r \circ g)(x)$, we see that

$$\begin{aligned}
(r \circ g)(x) &= r(x^2 + c) \\
&= p(x^2 + c) - q(x^2 + c) \\
&= (p^2(x) + c) - (q^2(x) + c) \\
&= p^2(x) + c - q^2(x) - c \\
&= p^2(x) - q^2(x) \\
&= (p(x) - q(x))(p(x) + q(x)) \\
&= r(x)(p(x) + q(x)).
\end{aligned}$$

The degree of the polynomial $(r \circ g)(x)$ is $2t$ by **2.9**. Also, the degree of the polynomial $r(x)(p(x)+q(x))$ is $t+k$, as both $p(x)$ and $q(x)$ are of degree k and both have a leading coefficient of 1. Then $2t = t + k$, or $t = k$, which is a contradiction. Therefore, $r(x)$ is the zero polynomial and $p(x) = q(x)$. Because assuming that there exist two distinct polynomials of degree k that commute with $g(x)$ led to a contradiction, we may conclude that there exists, at most, one polynomial of degree k that commutes with $s(x)$ under composition. $\square$

The above result and method of proof can be generalized to provide a much stronger result. This generalization is the main theorem of this section.

Theorem 3.3 *There are at most $n - 1$ polynomial(s) of degree $k \geq 1$ that commute with a given polynomial*

$$s(x) = \sum_{m=0}^{n} a_m x^m$$

with $a_n \neq 0$ and $n \geq 2$.

Proof. By **2.10**, we may define

$$\lambda(x)$$

of degree 1 such that we obtain

$$(\lambda^{-1} \circ s \circ \lambda)(x) = h(x),$$

where

$$h(x) = \sum_{m=0}^{n} b_m x^m, b_n = 1.$$

It suffices to show that there are, at most, $n-1$ polynomials of degree $k \geq 1$ that commute with $h(x)$, as $h(x)$ is similar to $s(x)$.

We know, by the **3.1**, that if a polynomial of any degree commutes with $h(x)$, then the leading coefficient of that polynomial must be a root ρ of $x^{n-1}-1$.

Assume, for the sake of contradiction, that there exist two polynomials $p(x)$ and $q(x)$ of degree $k \geq 1$ that both commute with $h(x)$, each with leading coefficient ρ.

Let $r(x) = (p-q)(x)$. We know that $\deg(r) = t \leq k-1$, since $p(x)$ and $q(x)$ have the same leading coefficient. We also note that

$$r(h(x)) = (p-q)(h(x)) = p(h(x)) - q(h(x))$$

$$= h(p(x)) - h(q(x))$$

$$= \sum_{m=1}^{n} b_m (p^m(x) - q^m(x))$$

by the setup, along with the fact that $p(x)$ and $q(x)$ commute with $h(x)$ under composition. When we factor out $p(x) - q(x)$ of the equation, we have

$$(p(x) - q(x)) \sum_{i=1}^{n} b_i \left(\sum_{j=0}^{i-1} p^{(i-1)-j}(x) q^j(x) \right).$$

Since $h(x)$ is monic, the leading expression of the segment that appears to the right of the first summation symbol is given by

$$\sum_{j=0}^{n-1} p^{(n-1)-j}(x)q^j(x)$$

(notice that the constant coefficients cancel each other out). The degree of the above summation is $(n-1)k$. This is because all summand polynomials determined by the summation notation are of degree $(n-1)k$, and the leading coefficient (that is, the coefficient of degree $(n-1)k$) of the summation is given by

$$\sum_{j=0}^{n-1} \rho^{(n-1)-j}\rho^j = \sum_{j=0}^{n-1} \rho^{n-1} = \sum_{j=0}^{n-1} 1 = n.$$

Because we know $p(x) - q(x) = r(x)$, we see that the degree of the entire polynomial $r(h(x)))$ is

$$\deg(r) + (n-1)k.$$

Then we have

$$\deg r(h(x)) = t + (n-1)k$$

$$nt = t + (n-1)k$$

$$nt - t = (n-1)k$$

$$t(n-1) = (n-1)k$$

$$t = k.$$

This is a contradiction, as $\deg(r(x)) \leq k-1$.

Therefore, we may conclude that, if $p(x)$ and $q(x)$ are both of degree k, they cannot share the same leading coefficient. Subsequently, we may conclude that for $k \geq 1$, there are (at most) as many polynomials of degree k that commute with $h(x)$ as there are roots of $x^{n-1} - 1$.

Because we are dealing with complex coefficients, we know that there are $n-1$ roots of $x^{n-1}-1$. So, there are at most $n-1$ polynomials of degree $k \geq 1$ that commute with $h(x)$. We defined $h(x)$ to be similar to $s(x)$. Therefore, we may finally conclude that there are, at most, $n-1$ polynomials of degree $k \geq 1$ that commute with a given polynomial of degree $n \geq 2$. $\square$

4 A Theorem of Block and Thielman

We have already mentioned two polynomial sequences: the Chebyshev polynomials and the power polynomials. These two sequences are extremely significant for two reasons. First, all of the polynomials in each sequence commute with every other polynomial in the sequence. Second, as we will soon see, that these polynomials play an extremely significant role in the characterization of other polynomial sequences.

Definition 4.1 *Let $\{f_j(x)|j = 1, 2, 3, \ldots\}$, be a sequence of polynomials that contains one polynomial of each positive degree. If every polynomial in $\{f_j(x)\}$ commutes with every other polynomial in the sequence, $\{f_j(x)\}$ is called a chain.*

The main result of this chapter characterizes chains of polynomials with complex coefficients. This result was first theorized and subsequently proven by Block and Thielman [2] in 1951, and later proven independently by Jacobsthal [3] in 1955. While we can easily find pairs of polynomials that commute with each other under composition, we can also readily find chains of polynomials. For instance, we can construct other chains, based on the two major chains already mentioned, via similarity. For example, using $\lambda(x) = x + 1$, we see that

$$(\lambda^{-1} \circ (x^j) \circ \lambda)(x), \quad j = 1, 2, 3, \ldots$$

is also a chain. We actually examined two polynomials in this chain earlier in this paper, and more may be evaluated. The above similarity converts the following power polynomials to new polynomials that, together, make up part of this new chain:

$$\begin{aligned} x &\to x \\ x^2 &\to x^2 + 2x \\ x^3 &\to x^3 + 3x^2 + 3x \\ x^4 &\to x^4 + 4x^3 + 6x^2 + 4x. \end{aligned}$$

A similar observation may be made by applying the similarity to the entire chain of Chebyshev polynomials. We can see that

$$(\lambda^{-1} \circ T_j \circ \lambda)(x), \quad j = 1, 2, 3, \ldots$$

is also a chain.

Both of these new chains are obviously similar to either the power polynomials or the Chebyshev polynomials. Their very construction shows this similarity. It turns out that the two original sequences are extremely significant, in that all chains of polynomials are similar to either the power polynomials or the Chebyshev polynomials. This is the main result of this chapter for which we will now set the stage.

Before we start the theorem and present its proof, we first recall a few definitions and establish a few lemmas. These first definitions recall two classes of polynomials from elementary calculus.

Definition 4.2 *Let $f(x)$ be a polynomial. We say that $f(x)$ is even if and only if $f(-x) = f(x)$. We say that $f(x)$ is odd if and only if $f(-x) = -f(x)$.*

The definitions of even polynomials and odd polynomials could be considered opposites. There is, however, one scenario that the above two definitions to not account for. Let $f(x)$ be a polynomial. If $f(-x) \neq f(x)$ and $f(-x) \neq -f(x)$, then $f(x)$ is neither even nor odd.

One interesting application of the above definition is that a polynomial $f(x)$ is odd if and only if it commutes with $g(x) = -x$. When we assume $f(x)$ is odd, we know that $f(-x) = -f(x)$, by definition. Thus, $f(g(x)) = g(f(x))$. Assuming that $f(x)$ and $g(x)$ commute implies that $f(g(x)) = g(f(x))$, which shows that $f(-x) = -f(x)$.

These definitions allow us to classify certain polynomials. The following lemmas rely on these classifications, and are necessary for the later proof of the Block Thielman theorem.

Lemma 4.3 *Let* $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$ *be an odd polynomial. Then, for all* j *such that* $j = 2k, \quad k = 0, 1, 2, \dots,$ $a_j = 0$.

Proof. Let $f(x)$ be as in the statement of the lemma. Because $f(x)$ is an odd polynomial, we have that

$$f(-x) = -f(x)$$

by definition. We know that

$$\begin{aligned} f(-x) &= a_0 + a_1(-x) + a_2(-x)^2 + a_3(-x)^3 + \dots \\ &= a_0 + (-a_1x) + a_2x^2 + (-a_3x^3) + \dots . \end{aligned}$$

We also know that

$$
\begin{aligned}
-f(x) &= -(a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots) \\
&= (-a_0) + (-a_1x) + (-a_2x^2) + (-a_3x^3) + \ldots.
\end{aligned}
$$

Recall that, in order for polynomials to be equal, the coefficients of like degrees must be equal. Therefore, for all j such that j is even, we have that

$$-a_j = a_j.$$

Therefore, $a_j = 0$, as desired. □

Lemma 4.4 *Let $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots$ be an even polynomial. Then, for all j such that $j = 2k+1, \quad j = 0, 1, 2, \ldots,$ $a_j = 0$.*

Proof. Let $f(x)$ be as in the statement of the lemma. Because $f(x)$ is an even polynomial, we have that

$$f(-x) = f(x)$$

by definition. We know that

$$
\begin{aligned}
f(-x) &= a_0 + a_1(-x) + a_2(-x)^2 + a_3(-x)^3 + \ldots \\
&= a_0 + (-a_1x) + a_2x^2 + (-a_3x^3) + \ldots.
\end{aligned}
$$

We also know that

$$f(x) \quad = \quad -(a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots).$$

Recall that, in order for polynomials to be equal, the coefficients of like degrees must be equal. Therefore, for all j such that j is odd, we have that

$$-a_j = a_j.$$

Therefore, $a_j = 0$, as desired. $\square$

We can now move on to the theorem of Block and Thielman, and Jacobstahl.

Theorem 4.5 *Every chain of polynomials is similar to (only) one of the two following chains:* $\{x^j, \quad j = 1, 2, 3, \ldots\}$ *(the power polynomials), and*$\{T_j(x), \quad j = 1, 2, 3, \ldots\}$ *(the Chebyshev polynomials).*

Proof. Let $\{p_j(x), \quad j = 1, 2, 3, \ldots\}$ be a chain, with $p_2(x) = a_2x^2 + a_1x + a_0$. Let $\{q_j(x), j = 1, 2, 3, \ldots\}$ be a chain similar to $\{p_j\}$ via $\lambda(x) = \dfrac{x}{a_2} - \dfrac{a_1}{2a_2}$. Then $q_2(x) = x^2 + c$, as shown earlier in the paper. We know that

$$q_3(x) = b_3x^3 + b_2x^2 + b_1x + b_0$$

commutes with $q_2(x)$ by the definition of chain. Thus, we have that

$$q_3(x^2 + c) = {q_3}^2(x) + c \quad (*)$$

and

$$q_3(x^2+c)-c = {q_3}^2(x)+c-c = {q_3}^2(x).$$

Therefore, we know that

$$\begin{aligned} {q_3}^2(-x) &= q_3((-x)^2+c)-c \\ &= q_3(x^2+c)-c \\ &= {q_3}^2(x), \end{aligned}$$

and it follows that

$$q_3(-x) = \pm q_3(x).$$

Therefore $q_3(-x) = -q_3(x)$ or $q_3(-x) = q_3(x)$. Assume, for the sake of contradiction, that $q_3(-x) = q_3(x)$. Then $q_3(x)$ is an even polynomial. By **4.4**, all of the coefficients of odd degrees in $q_3(x)$ are 0. This is a contradiction, as $\deg(q_3(x)) = 3$ and, by definition, $b_3 \neq 0$. Therefore, $q_3(-x) = -q_3(x)$ and $q_3(x)$ is an odd polynomial, by definition. By **4.3**, we may conclude that $q_3(x)$ is of the form

$$q_3(x) = b_3x^3 + b_1(x).$$

We know by **3.1** that $b_3 = 1$. So, when we substitute the above representation of $q_3(x)$ into equation $(*)$, we get

$$(x^2+c)^3 + b_1(x^2+c) = (x^3+b_1x)^2 + c$$

or

$$(x^6 + 3cx^4 + 3c^2x^2 + c^3) + b_1(x^2 + c) = x^6 + 2b_1x^4 + {b_1}^2x^2 + c.$$

This simplifies to

$$x^6 + 3cx^4 + (3c^2 + b_1)x^2 + c^3 + b_1c = x^6 + 2b_1x^4 + {b_1}^2x^2 + c.$$

Comparing the coefficients of x^4 shows that

$$3c = 2b_1,$$

so $b_1 = \frac{3}{2}c$. Knowing b_1, we move on to compare the coefficients of x^2. We see that

$$\begin{aligned}
3c^2 + b_1 &= (b_1)^2 \\
3c^2 + \frac{3}{2}c &= (\frac{3}{2}c)^2 \\
3c^2 + \frac{3}{2}c &= \frac{9}{4}c^2 \\
\frac{3}{4}c^2 + \frac{3}{2}c &= 0 \\
3c^2 + 6c &= 0 \\
c^2 + 2c &= 0 \\
c(c+2) &= 0.
\end{aligned}$$

Finally, when we compare the constant values in each equation, we see that

$$c^3 + \frac{3}{2}c^2 = c$$

$$
\begin{aligned}
2c^3 + 3c^2 &= 2c \\
2c^3 + 3c^2 - 2c &= 0 \\
(c+2)(2c^2 - c) &= 0 \\
c(c+2)(2c-1) &= 0.
\end{aligned}
$$

These last two coefficient comparisons show that $c = -2$ or $c = 0$.

Case 1: Assume $c = 0$. Then $q_2(x) = x^2$. By **3.2**, the only polynomials that commute with $q_2(x)$ are elements of $\{x^j\}$. Therefore, we know that $\{q_j(x)\} = \{x^j\}$ and that $\{p_j(x)\}$ is similar to $\{x^j\}$.

Case 2: Assume $c = -2$. Then $q_2(x) = x^2 - 2$. Now consider the degree 1 polynomial

$$\alpha(x) = 2x.$$

We can see that

$$
\begin{aligned}
(\alpha^{-1} \circ q_2 \circ \alpha)(x) &= \alpha^{-1}((2x)^2 - 2) \\
&= \alpha^{-1}(4x^2 - 2) \\
&= \frac{1}{2}(4x^2 - 2) \\
&= 2x^2 - 1 \\
&= T_2(x).
\end{aligned}
$$

Now, because every element of $\{q_j(x),\ j = 1, 2, 3, \ldots\}$ commutes with one another, we know that every element of $\{(\alpha^{-1} \circ q_j \circ \alpha)(x),\ j = 1, 2, 3, \ldots\}$, commutes with one another by **2.11**. Also, because $(\alpha^{-1} \circ q_2 \circ \alpha)(x) = T_2(x)$, we know that any polynomial that commutes with $(\alpha^{-1} \circ q_2 \circ \alpha)(x)$ must be an element of $\{T_j(x)\}$ by **3.2**. Therefore, $\{(\alpha^{-1} \circ q_j \circ \alpha)(x)\} = \{T_j(x)\}$. Because similarity is transitive, we know that $\{p_j(x)\}$ is also similar to $\{T_j(x)\}$.

Because the above cases are exhaustive, we conclude that $\{p_j(x)\}$ is either similar to $\{x^j\}$ or $\{T_j(x)\}$, as desired. $\square$

5 Self Similarity

For the final chapter of this paper, we will prove an interesting result that relates to the number of polynomials of a given degree that can commute with a given polynomial.

Definition 5.1 *Let $f(x)$ be a polynomial. If $f(x)$ is similar to itself via the similarity $\lambda(x)$, we say that $\lambda(x)$ is a self-similarity of $f(x)$. We will denote the set of self-similarities of $f(x)$ as S_f.*

For example, let $f(x) = x^3$ and consider $\lambda(x) = -x$. Then, we have that

$$(\lambda^{-1} \circ f \circ \lambda)(x) = \lambda^{-1}(f(-x)) = \lambda^{-1}((-x)^3) = \lambda^{-1}(-(x^3)) = x^3 = f(x).$$

Thus, $\lambda(x) = -x$ is a self-similarity of $f(x) = x^3$. We know $f(x) = x^3$ is an odd function by **4.3**, so $f(-x) = -f(x)$. Thus, $f(x)$ commutes with its self-similarity $\lambda(x) = -x$. This result is true in general.

Proposition 5.2 *Let $f(x)$ be a polynomial and let $\lambda(x)$ be a degree 1 polynomial. Then $\lambda(x) \in S_f$ if and only if $f(x)$ and $\lambda(x)$ commute under composition.*

Proof. Assume that $\lambda(x)$ is a self-similarity of $f(x)$. We know that

$$(\lambda^{-1} \circ f \circ \lambda)(x) = f(x).$$

When we compose each side of the equation with $\lambda(x)$, we get

$$(\lambda \circ \lambda^{-1} \circ f \circ \lambda)(x) = (\lambda \circ f)(x)$$

which simplifies to

$$(f \circ \lambda)(x) = (\lambda \circ f)(x).$$

Thus, we conclude that $f(x)$ and $\lambda(x)$ commute under composition, as desired.

The proof of the converse is similar, but opposite. We have that

$$(f \circ \lambda)(x) = (\lambda \circ f)(x).$$

Therefore, we can easily show that

$$(\lambda^{-1} \circ (f \circ \lambda))(x) = (\lambda^{-1} \circ (\lambda \circ f))(x)$$

which can be rewritten as

$$(\lambda^{-1} \circ f \circ \lambda)(x) = (\lambda^{-1} \circ \lambda \circ f)(x).$$

But this means that

$$(\lambda^{-1} \circ f \circ \lambda)(x) = f(x),$$

as desired. □

By the above proof, we can see that S_f contains all of the degree one polynomials that commute with $f(x)$ under composition.

Definition 5.3 *Let* $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$ *be a polynomial. We say that* $f(x)$ *is* centered *if* $a_n \neq 0$ *and* $a_{n-1} = 0$.

For clarity, $O(x^k)$ will denote any polynomial with degree less than or equal to k.

Proposition 5.4 *Let $f(x)$ be a polynomial and* $\deg(f(x)) \geq 1$*. Then $f(x)$ is similar to a monic, centered polynomial.*

Poof. Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

be a polynomial and let $f(x)$ be similar to $g(x)$ via the similarity

$$\lambda(x) = bx + c.$$

Then we have that

$$\begin{aligned}
g(x) &= (\lambda^{-1} \circ f \circ \lambda)(x) \\
&= \lambda^{-1}(f(\lambda(x))) \\
&= \lambda^{-1}(f(bx+c)) \\
&= \lambda^{-1}(a_n(bx+c)^n + a_{n-1}(bx+c)^{n-1} + \ldots + a_2(bx+c)^2 + a_1(bx+c) + a_0) \\
&= \lambda^{-1}(a_n b^n x^n + a_n(b^{n-1}nx^{n-1}c) + a_{n-1}b^{n-1}x^{n-1} + O(x^{n-2})) \\
&= \frac{(a_n b^n x^n + a_n(b^{n-1}nx^{n-1}c) + a_{n-1}b^{n-1}x^{n-1})}{b} + O(x^{n-2}) \\
&= \frac{(b^n a_n x^n + b^{n-1}(na_n c + a_{n-1})x^{n-1}}{b} + O(x^{n-2})) \\
&= b^{n-1}a_n x^n + b^{n-2}(na_n c + a_{n-1})x^{n-1} + O(x^{n-2}).
\end{aligned}$$

So, for $g(x)$ to be monic, let b be as it was in the proof of **2.10**. That is, let b be an $(n-1)^{st}$ root of ${a_n}^{-1}$. Then

$$b^{n-1}a_nx^n = {a_n}^{-1}a_nx^n = x^n.$$

For $g(x)$ to be centered, let $c = -\dfrac{a_{n-1}}{na_n}$. Then

$$na_nc + a_{n-1} = na_n(-\frac{a_{n-1}}{na_n}) + a_{n-1} = -a_{n-1} + a_{n-1} = 0.$$

□

We can now prove two results about polynomials that commute with a centered polynomial.

Proposition 5.5 *Let* $f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$ *be a polynomial with* $\deg(f(x)) = n > 1$ *and let* $p(x)$ *commute with* $f(x)$*, with* $\deg(p(x)) = k \geq 1$*. Assume* $f(x)$ *is monic and centered and* $p(x) = b_kx^k + b_{k-1}x^{k-1} + \ldots + b_2x^2 + b_1x + b_0$*. Then:*

(1) For some $j = 1, 2, \ldots, n-1$*,* $b_k = \rho^j$ *where* ρ *is a* $(n-1)^{\text{st}}$ *root of unity,*

(2) $b_{k-1} = 0$*.*

Proof. For the first result, we recall the proof of **3.1** and compare the leading coefficients of the polynomials $(f \circ p)(x) = (p \circ f)(x)$. It is easy to see that

$$a_n{b_k}^n = b_k{a_n}^k.$$

Since $f(x)$ is monic, we have that

$$b_k{}^n = b_k.$$

or

$$b_k{}^{n-1} = 1,$$

proving the first result. For the second result, we compare the the degree $kn - 1$ coefficients. We see that

$$kb_k a_n{}^{k-1} a_{n-1} = a_n b_k{}^{n-1} b_{k-1} n.$$

The previous result, combined with the fact that $f(x)$ is monic and centered, reveals that

$$a_n b_k{}^{n-1} = 1.$$

Because $f(x)$ is centered, we know that

$$a_{n-1} = 0.$$

Thus, we have

$$(0)(kb_k a_n{}^{k-1}) = b_{k-1} n(1)$$

or

$$0 = b_{k-1},$$

since $n \neq 0$, as desired. $\square$

Corollary 5.6 *Let $f(x)$ be a polynomial,* $\deg(f(x)) = n > 1$ *and let* $p(x) \neq x$ *commute with* $f(x)$ *with* $\deg(p(x)) = k \geq 1$. *Then $f(x)$ is centered if and only if $p(x)$ is centered.*

Proof. Assume $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x_1 + a_0$ is centered, so $a_{n-1} = 0$. We know that $f(x)$, via a similarity $\lambda(x)$, is similar to a monic polynomial by the earlier result shown in **5.4**. Let $\lambda(x) = bx$, with b defined as it is in the proof of **5.4**. Then $(\lambda^{-1} \circ f \circ \lambda)(x)$ is monic. Also, the construction of $\lambda(x)$ does not affect centering, so $(\lambda^{-1} \circ f \circ \lambda)(x)$ is also centered.

Since $p(x)$ commutes with $f(x)$, we know that $(\lambda^{-1} \circ p \circ \lambda)(x)$ also commutes with $(\lambda^{-1} \circ f \circ \lambda)(x)$ by **2.11**. Then, by **5.5**, $(\lambda^{-1} \circ p \circ \lambda)(x)$ is centered. Because the similarity does not affect centering, we conclude that $p(x)$ is also centered, as desired.

A similar argument may be used to prove the converse unless $\deg(p(x)) = 1$. Specifically, if $p(x) = cx$, $c \neq 0, 1$, then proposition **5.5** doesn't apply.

Without loss of generality, we can set

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x_1 + a_0.$$

We know $a_n \neq 0$ because $\deg(f(x)) = n$. Assume, for contradiction's sake, that $a_{n-1} \neq 0$. Then we have that

$$f(p(x)) \;=\; f(cx)$$

$$
\begin{aligned}
&= a_n(cx)^n + a_{n-1}(cx)^{n-1} + \ldots + a_2(cx)^2 + a_1(cx) + a_0 \\
&= a_n c^n x^n + a_{n-1}c^{n-1}x^{n-1} + \ldots + a_2c^2x^2 + a_1cx + a_0(*).
\end{aligned}
$$

However, we also know that $f(x)$ and $p(x)$ commute under composition, so

$$
\begin{aligned}
f(p(x)) &= p(f(x)) \\
&= c * (a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x_1 + a_0) \\
&= a_n c x^n + a_{n-1}cx^{n-1} + \ldots + a_2cx^2 + a_1cx + a_0.
\end{aligned}
$$

If we compare coefficients of like degrees between the above equation and equation (*), we see that

$$a_n c = a_n c^n$$

and

$$a_{n-1}c = a_{n-1}c^{n-1}.$$

These simplify to

$$c^n = c = c^{n-1},$$

which implies that $c = 0$ or $c = 1$, a contradiction. Therefore, we conclude that $a_{n-1} = 0$ and that $f(x)$ is centered, as desired. □

Before approaching the main result of the chapter, we will define a function, and note that this function is a bijection.

Lemma 5.7 *Assume that $f(x)$ and $g(x)$ are polynomials, and that $f(x)$ is similar to $g(x)$ via $\lambda(x)$. Let $\alpha(x) \in S_f$, and define $\phi(\alpha(x)) = (\lambda^{-1} \circ \alpha \circ$*

$\lambda)(x)$. *Then* $\phi : S_f \to S_g$ *is a bijection.*

Proof. Let $f(x)$, $g(x)$, $\lambda(x)$, and $\phi(\alpha(x))$ be as in the statement of the lemma. We know that, for all $\alpha(x) \in S_f$, $\alpha(x)$ commutes with $f(x)$ by **5.2**. Then $\phi(\alpha(x)) = (\lambda^{-1} \circ \alpha \circ \lambda)(x)$ commutes with $(\lambda^{-1} \circ f \circ \lambda)(x) = g(x)$ by **2.11**. That is, all of the elements of the codomain commute with $g(x)$. The similarity does not affect degree, so the elements of the codomain are also degree one polynomials, and we conclude that the codomain is S_g by **5.2**.

To show that ϕ is bijective, we will show that it is injective and surjective. To show that ϕ is injective, take any $\alpha(x)$, $\beta(x) \in S_f$. If $\phi(\alpha(x)) = \phi(\beta(x))$, we have that $(\lambda^{-1} \circ \alpha \circ \lambda)(x) = (\lambda^{-1} \circ \beta \circ \lambda)(x)$, and composing the left side and right side of each equation with $\lambda(x)$ and $\lambda^{-1}(x)$, respectively, shows that $\alpha(x) = \beta(x)$. Thus, $\phi(\alpha(x))$ is injective.

To show that ϕ is surjective, take any $\beta(x) \in S_g$ and consider $\alpha(x) = (\lambda \circ \beta \circ \lambda^{-1})(x)$. Then $\phi(\alpha(x)) = (\lambda^{-1} \circ \lambda \circ \beta \circ \lambda^{-1} \circ \lambda)(x) = \beta(x)$, as desired.

Because ϕ is both injective and surjective, we conclude that ϕ is bijective, as desired. □

Corollary 5.8 *Let* $f(x)$ *be a monic, centered polynomial and* $\deg(f(x)) = n > 1$. *Then any self-similarity* $\lambda(x)$ *of* $f(x)$ *is of the form* $\lambda(x) = p^j x$, *where* p *is a primitive* $(n-1)^{st}$ *root of unity.*

Proof. We know that $f(x)$ commutes with $\lambda(x)$. Therefore, by **5.5**, the result immediately follows. □

Corollary 5.9 *Let $f(x)$ be a polynomial and* $\deg(f(x)) = n > 1$*. Then the order of S_f is less than or equal to $n-1$.*

Proof. By **5.4**, we know that $f(x)$ is similar to a monic, centered polynomial $g(x)$. We know that the order of S_g is less than or equal to $n-1$ because, by **5.8**, all of the elements of S_g are of the form $\lambda(x) = p^j x$, where p is a primitive $(n-1)^{st}$ root of unity. Therefore, there are only $n-1$ possible unique elements in S_g.

However, we also know that the order of S_f is equal to the order of S_g by lemma **5.7** since ϕ is bijective. Therefore, we conclude that the order of S_f is less than or equal to $n-1$, as desired. □

We may now move on to proving the main result of this chapter. The proof of the following theorem, originally given by Zimmermann [5], is similar to the proof of **3.3**.

Theorem 5.10 *Let $f(x)$ be a polynomial and* $\deg(f(x)) = n > 1$*. Also assume $p(x)$ and $q(x)$ commute with $f(x)$ under composition and* $\deg(p(x)) = \deg(q(x)) = k \geq 1$*. Then $q(x) = (\lambda_f \circ p)(x)$, where $\lambda_f(x)$ is a self-similarity of $f(x)$.*

Proof. By **5.4**, we know that $f(x)$ is similar to a monic, centered polynomial

$$g(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$$

via some similarity $\lambda(x)$.

Now, let

$$P(x) = (\lambda^{-1} \circ p \circ \lambda)(x)$$

and

$$Q(x) = (\lambda^{-1} \circ q \circ \lambda)(x).$$

By **5.5**, it follows that

$$P(x) = d^i x^k + b_{k-2}x^{k-2} + \ldots + b_2x^2 + b_1x + b_0$$

and

$$Q(x) = d^j x^k + h_{k-2}x^{k-2} + \ldots + h_2x^2 + h_1x + h_0,$$

where d is a primitive $n - 1^{st}$ root of unity. Let $\alpha = j - i$.

Now let $r(x) = d^\alpha P(x) - Q(x)$. and note that $\deg(r(x)) = t < k$ or $r(x) = 0$, since the leading coefficients of $d^\alpha P(x)$ and $Q(x)$ agree. We have that

$$r(g(x)) = d^\alpha P(g(x)) - Q(g(x)) = d^\alpha g(P(x)) - g(Q(x)),$$

since we know that $P(x)$ and $Q(x)$ commute with $g(x)$ under composition by **2.11**. Also, because $g(x)$ is monic and centered, we know that $a_n = 1$ and $a_{n-1} = 0$. Note that, because d is a $n - 1^{st}$ root of unity, we have that $d^\alpha = 1^\alpha d^\alpha = (d^\alpha)^{n-1}(d^\alpha)^1 = (d^\alpha)^n$.Then we have that

$$r(g(x)) \quad = \quad d^\alpha g(P(x)) - g(Q(x))$$

$$
\begin{aligned}
&= d^{\alpha} \sum_{m=0}^{n} a_m (P(x))^m - \sum_{m=0}^{n} a_m (Q(x))^m \\
&= d^{\alpha}(P(x))^n - (Q(x))^n) + O(x^{k(n-2)}) \\
&= (d^{\alpha})^n (P(x))^n - (Q(x))^n) + O(x^{k(n-2)}) \\
&= (d^{\alpha}P(x))^n - Q^n(x) + O(x^{k(n-2)}) \\
&= (d^{\alpha}P(x) - Q(x)) \left(\sum_{s=0}^{n-1} (d^{\alpha}P(x))^{n-1-s} (Q(x))^s \right) + O(x^{k(n-2)}) \\
&= r(x) \left(\sum_{s=0}^{n-1} (d^{\alpha}P(x))^{n-1-s} (Q(x))^s \right) + O(x^{k(n-2)}) \quad (*)
\end{aligned}
$$

Note that $\deg \left[\sum_{s=0}^{n-1} (d^{\alpha}P(x))^{n-1-s} (Q(x))^s \right] = k(n-1)$ since the leading coefficient, that is, the degree $k(n-1)$ coefficient, is

$$
\sum_{s=0}^{n-1} (d^{\alpha})^{n-1-s} (d^i)^{n-1-s} (d^j)^s = \sum_{s=0}^{n-1} (d^j)^{n-1-s} (d^j)^s = \sum_{s=0}^{n-1} (d^{(n-1)})^j (d^{-s})^j (d^j)^s =
$$

$$
\sum_{s=0}^{n-1} 1 = n \neq 0
$$

because d is a primitive $n-1^{st}$ root of unity.

Comparing degrees in the two sides of the equation for $r(g(x))$, we see that $tn = \deg(r(g(x))) = t + k(n-1)$, which implies that $t = k$. This is a contradiction since $t < k$, and therefore $r(x) = 0$. Because $r(x) = 0$, we know that $0 = d^{\alpha}P(x) - Q(x)$, or that $Q(x) = d^{\alpha}P(x)$, as desired.

Consider $\lambda_g(x) = d^\alpha x$. To see that $\lambda_g(x)$ is a self-similarity of $g(x)$, observe that $g(Q(x)) = Q(g(x))$ implies

$$g(d^\alpha P(x)) = d^\alpha P(g(x)) = d^\alpha g(P(x)).$$

Letting $P(x) = u$, we have $g(d^\alpha u) = d^\alpha g(u)$. Thus, by **5.2**, $\lambda_g(x)$ is a self-similarity of $g(x)$, as desired.

We define $\lambda_f(x) = (\lambda \circ \lambda_g \circ \lambda^{-1})(x) = \phi^{-1}(\lambda_g(x))$, where ϕ is the function described in **5.7** and $\lambda(x)$ is the similarity introduced at the beginning of this proof. Thus, we may conclude that$\lambda_f(x)$ is a self-similarity of $f(x)$. It follows that

$$\begin{aligned} Q(x) &= (\lambda_g \circ P)(x) \\ (\lambda \circ Q \circ \lambda^{-1})(x) &= (\lambda \circ \lambda_g \circ P \circ \lambda^{-1})(x) \\ (\lambda \circ \lambda^{-1} \circ q \circ \lambda \circ \lambda^{-1})(x) &= (\lambda \circ \lambda^{-1} \circ \lambda_f \circ \lambda \circ \lambda^{-1} \circ p \circ \lambda \circ \lambda^{-1})(x) \\ q(x) &= (\lambda_f \circ p)(x) \end{aligned}$$

as desired. □

The degree analysis performed on equation $(*)$ is difficult, and it is therefore useful to evaluate the equation with $g(x)$ as a degree 4 polynomial. Then we have

$$r(x)\left(\sum_{s=0}^{3}(d^\alpha P(x))^{4-1-s}(Q(x))^s\right) + O(x^{2k})$$

which can be expanded to

$$(d^\alpha P(x) - Q(x))\Big((d^\alpha P(x))^3 + (d^\alpha P(x))^2(Q(x)) + (d^\alpha P(x))(Q(x))^2 + (Q(x))^3\Big)$$

$$+O(x^{2k}).$$

The degree of the polynomial within the large parentheses is $3k$, as the degree of each summand polynomial is $3k$ and the leading coefficient of the large polynomial is given by

$$(d^\alpha d^i)^3 + (d^\alpha d^i)^2 d^j + (d^\alpha d^i)(d^j)^2 + (d^j)^3 = d^{3(j-i)}d^{3i} + d^{2(j-i)}d^{2i}d^j + d^{j-i}d^i d^{2j} + d^{3j}$$

$$= d^{3j} + d^{3j} + d^{3j} + d^{3j} = 1 + 1 + 1 + 1 = 4.$$

Thus, the degree of the entire polynomial is $t+3k$, which is actually $t+k(n-1)$, illustrating the result we derived above.

The upcoming lemma provides the final results necessary for proving this chapter's most important result. Before we state the lemma and present the proof, however, we must recall some simpler results regarding polynomials.

Theorem 5.11 *Let $f(x)$ be a polynomial. Then a is a root of $f(x)$ if and only if $x - a$ is a factor of $f(x)$.*

Proof. Assume a is a root of $f(x)$. Then $f(a) = 0$. By the division algorithm, we have that

$$f(x) = (x - a)q(x) + r(x),$$

where $\deg(r(x)) < \deg(x-a) = 1$ or $r(x) = 0$. Assume, for the sake of contradiction, that $r(x) \neq 0$. Then $r(x)$ is a constant $c \neq 0$, and $f(a) = (x-a)(q(x))+c = c$. This is a contradiction, so we know that $r(x) = 0$. Thus, $f(x) = (x-a)(q(x))$, and we conclude that $x-a$ is a factor of $f(x)$, as desired.

Now assume that $x-a$ is a factor of $f(x)$, and $deg(f(x)) = n$. Then, without loss of generality, we may represent $f(x)$ as

$$f(x) = (x-a)O(x^{n-1}).$$

For $f(a)$, we have

$$f(a) = (a-a)O(x^{n-1}) = 0 * O(x^{n-1}) = 0,$$

and a is a root of $f(x)$, as desired. □

Lemma 5.12 *Let f(x) be a polynomial of degree $n > 0$. Then $f(x)$ has, at most, n roots.*

Proof. Assume, for the sake of contradiction, that $f(x)$ has more than n roots. Without loss of generality, assume $n+1$ roots. Then, by the above theorem, $f(x)$ could be represented as

$$f(x) = (x-a_{n+1})(x-a_n)(x-a_{n-1})\dots(x-a_2)(x-a_1)g(x)$$

where $a_{n+1}, a_n, a_{n-1}, \dots \in \mathbf{C}$ are all roots of $f(x)$. But, when these factors multiply together, they yield a leading coefficient of 1, corresponding to x^{n+1}.

Then $\deg(f(x)) = n + 1 + \deg(g(x))$, which is a contradiction. Thus, we may conclude that $f(x)$ has, at most, n roots, as desired. $\square$

Lemma 5.13 *Let $f(x)$ and $g(x)$ be polynomials, each with degree less than or equal to n. If $f(x)$ and $g(x)$ are equal when for $x = a_1, \ldots, a_{n+1}$, then $f(x) = g(x)$*

Proof. Assume, for the sake of contradiction, that $f(x)$ and $g(x)$ agree on $n+1$ distinct values and aren't equal. Consider $r(x) = (f - g)(x)$. We know that $\deg(r(x)) \leq n$, and we also know that $r(x)$ has $n + 1$ roots, since $f(x) = g(x)$ for $n + 1$ distinct points. This is a contradiction, as $r(x)$ cannot have more than n roots by the above lemma.
Therefore, we conclude that $f(x) = g(x)$, as desired. $\square$

We can now prove the lemma that leads to the main result of this chapter.

Lemma 5.14 *Let $f(x)$ and $\lambda(x)$ be polynomials, $\deg(f(x)) > 1$ and $\deg(\lambda(x)) = 1$. Assume $p(x)$ commutes with $f(x)$ and $\deg(p(x)) = k \geq 1$. Then $(\lambda \circ p)(x)$ commutes with $f(x)$ if and only if $\lambda(x) \in S_f$.*

Proof. Assume $(\lambda \circ p)(x)$ commutes with $f(x)$. Then we have

$$((\lambda \circ p) \circ f)(x) = (f \circ (\lambda \circ p))(x)$$

or

$$(\lambda \circ p \circ f)(x) = (f \circ \lambda \circ p)(x).$$

Composing $\lambda^{-1}(x)$ with each side of the equation yields

$$(p \circ f)(x) = (\lambda^{-1} \circ f \circ \lambda \circ p)(x).$$

Because $p(x)$ commutes with $f(x)$ under composition, we see that

$$(f \circ p)(x) = (\lambda^{-1} \circ f \circ \lambda \circ p)(x)$$

or

$$(f \circ p)(x) = ((\lambda^{-1} \circ f \circ \lambda) \circ p)(x)$$

so

$$f(p(x)) = (\lambda^{-1} \circ f \circ \lambda)(p(x))$$

Because $p(x)$ is non-constant, we can see that $f(x)$ and $(\lambda^{-1} \circ f \circ \lambda)(x)$ are equal when x is evaluated for an infinite number of values. Then, by lemma **5.13**, we know $f(x) = (\lambda^{-1} \circ f \circ \lambda)(x)$, and $\lambda(x) \in S_f$, as desired.

To prove the converse, assume $\lambda(x) \in S_f$. Then we have that

$$\begin{aligned} (f \circ \lambda \circ p)(x) &= (\lambda \circ \lambda^{-1} \circ f \circ \lambda \circ p)(x) \\ &= (\lambda \circ f \circ p)(x) \\ &= (\lambda \circ p \circ f)(x) \end{aligned}$$

as desired. □

We now prove the most important result of this chapter.

Theorem 5.15 *Let $f(x)$ be a polynomial with* $\deg(f(x)) = n > 1$ *and let $C_k(f(x))$ denote the set of polynomials of degree $k \geq 1$ that commute with $f(x)$ under composition. Assume $C_k(f(x))$ is non-empty for some k. Then the number of elements in $C_k(f(x))$ is the order of S_f.*

Proof. Let $p(x) \in C_k(f)$, and define $\psi : S_f \to C_k(f(x))$ by $\psi(\lambda) = \lambda \circ p$. By **5.14**, $\psi(\lambda(x)) \in C_k(f(x))$. We know $p(x)$ is non-constant, as it is a polynomial of degree greater than or equal to 1. Therefore, it follows that ψ is injective, as for any $\lambda_1(x)$, $\lambda_2(x) \in S_f$, if $\psi(\lambda_1(x)) = \psi(\lambda_2(x))$, we know that $\lambda_1(p(x)) = \lambda_2(p(x))$. Thus, by **5.13**, $\lambda_1(x) = \lambda_2(x)$.

Also, by **5.9**, we know that ψ is surjective, as all elements of $C_k(f(x))$ may be represented as the composition of $p(x)$ and a self-similarity of $f(x)$. Thus, we know that ψ is a bijection. Because ψ is a bijection, we may conclude that the order of $C_k(f(x))$ is the order of S_f, as desired. □

The above result provides a proof for a result already shown in this paper.

Corollary 5.16 *Let $f(x)$ be a polynomial and* $\deg(f(x)) = n > 1$*. Then there are at most $n - 1$ polynomials of degree $k \geq 1$ that commute with $f(x)$ under composition.*

Proof. This result follows immediately from **5.14**, indicating that the order of $C_k(f(x))$ is equal to the order of S_f, and **5.8**, which shows that the order of S_f is less than or equal to $n - 1$. □

References

[1] E. J. Barbeau, Polynomials, Springer-Berlag, 1989, pp 16.

[2] H. D. Block and H. P. Thielman, Commutative polynomials, Quart. J. Math. Oxford Ser., (2), 2 (1951) 241-243.

[3] E. J. Jacobsthal, Über vertauschbare polynome, Math. Z., 63 (1955) 243-276.

[4] T. J. Rivlin, Chebyshev Polynomials, Wiley Interscience, 1971, pp 194.

[5] K. Zimmermann, Commuting polynomials and self-similarity, New York J. Math., 13 (2007) 91-94.